The Report Committee for Andrias Rebecca Jones
Certifies that this is the approved version of the following Report:

Deliberative Groups and Their Decisions: An Application to State Revenue Estimation Practices

APPROVED BY
SUPERVISING COMMITTEE:

Martin Luby, Supervisor

Sean Williams, Co-Supervisor

Deliberative Groups and Their Decisions:
An Application to State Revenue Estimation Practices

by

Andrias Rebecca Jones

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degrees of

Master of Public Affairs

and

Doctor of Jurisprudence

The University of Texas at Austin

May 2018

Abstract

Deliberative Groups and Their Decisions:
An Application to State Revenue Estimation Practices

Andrias Rebecca Jones, M.P.Aff., J.D.

The University of Texas at Austin, 2018

Supervisor: Martin Luby and Sean Williams

States use a variety of methods to forecast revenue estimates, from executive branch agencies to consensus groups incorporating representatives from multiple branches of government. Findings from the behavioral economics literature suggest that deliberative groups systematically produce relatively biased and inaccurate outputs. The literature also describes characteristics of deliberative groups most likely to perform best, finding that small groups and majoritarian groups perform better at estimative tasks requiring high levels of accuracy. This paper applies these predictions in a revenue estimation context, comparing between consensus and non-consensus states and within consensus states. Findings confirm hypotheses drawn from the behavioral economics literature: consensus groups performed less accurately on average than non-consensus groups; unanimous groups generated less accurate forecasts than groups operating under majoritarian rules; and small consensus groups produced more accurate forecasts than big consensus groups.

Table of Contents

List of Tables

List of Figures

Introduction

Accurate revenue predictions rely on predicting the future of the economy, an almost impossible task. States use a variety of methods in attempting to forecast a correct future revenue projection. Texas, for instance, relies the Office of the Comptroller, a state agency somewhat unusually headed by an official who wins statewide election, to produce a biennial revenue forecast binding on legislative budget writers.[1] Other states rely on projections produced by expert-staffed agencies not subject to elections, governors' offices, or state legislatures.[2] One widespread method of projecting future state revenues is consensual revenue estimation, in which representatives from multiple branches of government or multiple agencies agree on a single forecast. Some states convene an independent board or commission, in which a group of political and expert non-political members determine revenue projections.

The accuracy of a group's decisionmaking process is crucial in this endeavor. When revenue estimators misfire, state legislators risk overspending beyond their budgets, or cutting unnecessary corners in state services. Pinpointing errors in existing revenue estimation practices and suggesting potentially superior alternatives could strengthen state-level governance and improve citizens' quality of life.

---

1 TEX. CONST. Art. III, Sec. 49a.

2 NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS, BUDGET PROCESSES IN THE STATES 31-38 (Spring 2015), https://higherlogicdownload.s3.amazonaws.com/NASBO/9d2d2db1-c943-4f1b-b750-0fca152d64c2/UploadedImages/Reports/2015%20Budget%20Processes%20-%20S.pdf.

However, behavioral economists studying the accuracy of decisions made by groups find that deliberative groups frequently make systematic mistakes. Deliberative groups fall prey to common biases that prevent members from attaining and correctly analyzing the best information. As a result, empirical studies show that these groups have inferior levels of factual accuracy compared to other methods of generating factual determinations.

Multimember groups, boards, and commissions responsible for estimating revenue operate as deliberative groups, and therefore may too experience group decisionmaking biases observed by behavioral economists. This possibility raises several questions. What testable hypotheses does behavioral economics generate about the accuracy of these groups' forecasts? What kinds of institutional arrangements produce the most accurate group forecasts? What alternatives to deliberative groups could states use to generate revenue forecasts of superior accuracy?

This paper will examine these questions. The first section gives an overview of the mechanics of state revenue estimation, as well as an introduction to some of the political, economic, and institutional determinants of forecast error. The second section discusses deliberative groups in the behavioral economics literature, and in particular the ways in which these groups amplify preexisting bias patterns. The third section generates a set of testable hypotheses regarding consensus budget estimation groups, a kind of deliberative group. Finally, the last section derives some workable suggestions for leaders managing forecasting processes.

I. State Revenue Estimation Processes

States accomplish revenue estimation using a large variety of different entities operating under different decisional rules and calendars. Each variable of the process can influence the quality of the final product, the revenue estimate. Different arrangements influence states' performance along several metrics, including not just the accuracy of the forecast but also levels of transparency and political acceptance.

A. Overview of the Mechanics of State Revenue Estimation

The internal process of developing the forecast vary between states. Most scholars divide states into a taxonomy with three categories: (1) separate forecasts by more than one branch of government, (2) executive forecasts by one or more executive branch agencies, or (3) consensus forecasts requiring multiple parts of government to come to an agreement on a single estimate.[3] A 2014 report estimates that 13 states employ separate forecasts, 10 states use executive-branch forecasts, and 28 states operate under some form of consensus forecasting.[4]

Scholars differ on the precise definition of "consensus." The National Association of State Budget Officers (NASBO) defines this type of forecast as a projection "developed in agreement through an official forecasting group representing both the executive and legislative branches. [5] Other definitions include not only representatives

---

3 Elizabeth C. McNichol, *Improving State Revenue Forecasting: Best Practices for a More Trusted and Reliable Revenue Estimate*, CENTER. ON BUDGET AND POLICY PRIORITIES 1-3 (Sep. 4, 2014) https://www.cbpp.org/sites/default/files/atoms/files/8-7-14sfp.pdf.

4 *Id.* at 9-10.

5 Donald J. Boyd & Lucy Dadayan, *State Tax Revenue Forecasting Accuracy Technical Report*, THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT 34 (Sep. 2014).

from executive and legislative branches but also experts and researchers from outside government. The Federation of Tax Administrators distinguishes between two types of consensus forecast groups: one type consisting of representatives of various state agencies, and a second type made up of executive and legislative appointees.[6]

The annual *Budget Processes in the States* almanac compiled by NASBO contains more detailed descriptions of the entities performing revenue estimation. They survey the fifty states to determine the entities responsible for preparing and revising revenue estimates. They categorize states by whether they use one or more of the following to estimate revenue: a budget or revenue agency, the governor's office, the state legislature, a board or commission, or some other type of entity, like as a governor-appointed state economist or group of university consultants.[7] According to NASBO, approximately 21 states also formally involved outside councils of economic advisors in their forecasting process.[8]

B. Determinants of Forecasting Error

Most studies show that state forecasters systematically underestimate state revenue.[9] Evidence suggests that this error rate is due to conscious decisions on the part of estimators rather than technical shortcomings of economic models.[10] Underestimation is a rational strategy for several possible reasons: as a hedge against economic

---

6 *Id.*

7 NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS, *supra* note 2, at 31-38.

8 *Id.* at 32.

9 Daniel W. Williams & Thad D. Calabrese, *The Status of Budget Forecasting*, 2 J. OF PUB. AND NONPROFIT AFFAIRS 127, 130 (2016); *see also* Boyd and Dadayan, *supra* note 5, at 14.

10 Williams & Calabrese, *supra* note 9, at 130.

uncertainty in the future; an ideologically motivated constraint on state spending; creating a stabilization cushion to forestall against future tax increases or service cuts; and so on.[11] In certain circumstances, underestimates may have pernicious consequences, especially if legislators choose to use their surplus on tax cuts or extra programs the state cannot afford in harder years.[12] Oregon has an unusual kicker law requiring state lawmakers to return revenues to taxpayers above 2 percent of the forecast; the kicker law forced the state to refund taxpayers $1.1 billion from a previous strong fiscal cycle while simultaneously issuing budget cuts of $1.3 billion.[13]

The precise magnitude and frequency of this underestimation is difficult to pin down. Previous studies have found that states underestimate on average approximately one to two percent of their revenue per fiscal cycle.[14] A recent comprehensive review of twenty-seven years' of revenue forecasting found that only five states regularly forecasted more than they collected; the other 45 states had median forecasts with positive errors.[15] States not only give low estimates of their overall tax collections, they also underestimate individual components of their total taxes including sales tax revenue, personal income tax revenue, and corporate income tax revenue.[16] US States share a propensity to underestimate revenue with sub-national units in other countries like

---

11 *Id.* at 130-31.
12 *Id.* at 13-14.
13 *Id.* at 15.
14 Shanna Rose & Daniel L. Smith, *Budget Slack, Institutions and Transparency*, 72 PUB. ADMIN. R. 187, 188 (2011).
15 Boyd & Dadayan, *supra* note 5, at 10-11.
16 *Id.* at 8.

Switzerland and Canada, but differ from national governments in the United States and Europe, which more commonly forecast more revenue than collected.[17]

States do at times overestimate revenues, usually in unusual circumstances like recessions. Estimators have an asymmetric incentive to avoid overestimation in anticipation of especially negative consequences: because many state budgets must balance, overestimates can force lawmakers to make targeted or across-the-board cuts, raise taxes, or spend down fiscal reserves.[18]

In the past two decades, both positive and negative forecasting errors have increased, perhaps as a reflection of increasing volatility in a less predictable economy. In a 23-year time span from 1987-2009, the median state estimate error fell within 1 percentage point only for two years.[19] From 1987-2001, these errors were more modest, falling within four percentage points of the mark. However, the median error increased over four points for five of the eight years from 2002 to 2009.[20] In 2009, the height of the recession, the median forecast was an extraordinary 10.9 percentage points over actual revenues, meaning half of all states performed even worse – in fact, four states returned an error greater than 25 percent.[21] As expected, errors increase during recessionary years,

---

17 Williams & Calabrese, *supra* note 9, at 131 and 143.

18 *Id.*

19 THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *States' Revenue Estimating: Cracks in the Crystal Ball* 17 (March 2011), http://www.rockinst.org/pdf/government_finance/2011-03-01-states-revenue-estimating-report.pdf.

20 *Id.*

21 *Id.* at 4.

but also during recent business expansion cycles.[22] From 2002-2006 median errors showed great volatility, fluctuating 6 to 8 percentage points off the mark. Even from 2010-2013, a period of steadier growth, median error stubbornly remained higher than 2 percentage points.[23] The next sections will explore factors demonstrated by the literature to affect revenue forecast errors.

1. Political Factors

A state's political context can and does influence revenue estimates. Boylan (2008) found that revenue estimators overestimated state revenues by 1.2-2.2 percentage points during gubernatorial election years, in particular when the incumbent's party faced serious threat.[24] While most scholars studying forecast accuracy include controls for unified government and the party in power, the evidence that these factors seriously influence accuracy is mixed.[25]

2. Economic Factors and Forecasting Methodology

Forecast accuracy is influenced by the economic context as well as forecasters' ability to interpret that context accurately. At the broadest level, the national economic climate affects state economies, and consequently state tax collections. While revenue estimators have always struggled with predicting recessions, economic volatility and

---

22 *Id.* at 11-12.

23 Boyd & Dadayan, *supra* note 5, at 8.

24 Richard T. Boylan, *Political Distortions in State Forecasts,* 136 PUB. CHOICE 411, 419-20 (2008); *see also* George A. Krause & James W. Douglas, *Organizational Structure and the Optimal Design of Policymaking Panels: Evidence from Consensus Group Commissions' Revenue Forecasts in the American States,* 57 AM. J. OF POL. SCI. 135, 143 (2013).

25 Rose & Smith, *supra* note 14, at 190.

unpredictability have increased progressively during each of the last three recessionary cycles.[26] Generally, most studies of state revenue forecasting control for macroeconomic variables including changes in unemployment, inflation, and personal income growth on a state and national level.[27]

Most forecasters use a three-step process: first, projecting nationwide economic trends; second, developing statewide economic projections with a model incorporating nationwide projections; and third, feeding state economic projections into econometric models of the state's tax revenues, taking into account the tax base and laws.[28] Estimators use a host of methodological tools, both quantitative and non-quantitative, to extrapolate into the future, including time series, causal and econometric modeling, simulations, Delphi methods, nominal techniques, and their own judgments.[29] However, even highly sophisticated modeling and tools sometimes do not generate additional explanatory power: "[I]mportant parts of the forecasting literature question the value added by attempting complex causal modeling of the variable to be forecast and propose that naïve

---

26 John L. Mikesell & Justin M. Ross, *State Revenue Forecasts and Political Acceptance: The Value of Consensus Forecasting in the Budget Process*, 74 PUB. ADMIN. REV. 188, 189 (2014); THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *supra* note 19, at 3.

27 *See, e.g.*, Boylan, supra note 23, at 417; Krause & Douglas, *supra* note 23, at 143; George A. Krause, David E. Lewis, & James W. Douglas, *Political Appointments, Civil Service Systems, and Bureaucratic Competence: Organizational Balancing and Executive Branch Revenue Forecasts in the American States*, 50 AM. J. POL. SCI. 770, 781 (2006).

28 Mikesell & Ross, *supra* note 25, at 189; THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *supra* note 19, at 17.

29 Williams & Calabrese, *supra* note 9, at 131 and 136-39.

models do just as well, if not better, and are less prone to interference with the forecast result."[30]

The structure of state tax laws and the anticipated mix of tax revenue seriously increases the difficulty of revenue forecasters' jobs. Frequently, revenue estimators must forecast multiple revenue streams for each source of tax revenue. [31] The three most important sources of revenue come from personal income tax, sales tax, and corporate income taxes, which together constitute on average 72 percent of a state's budget; a 2014 analysis showed that other sources of revenue made up more than half of a state's total revenue in only six states.[32] Each revenue source can inject volatility or unpredictability into the process. In particular, states in which income, sales, and corporate taxes constitute a minority of total revenue experience greater volatility, as do states with an over-reliance on a corporate tax.[33]

3. Institutional Factors: Generally

The rules and laws of the budget process have a significant effect on the forecasting accuracy. The nature of the state's fiscal rules and laws matter, as do the nature of unwritten rules and norms, as so many revenue estimating committees operate on an informal basis.

Several institutional factors influence the accuracy of a revenue forecast. A variety of tax and expenditure limits tie appropriators' hands and place additional

---

30 Mikesell & Ross, *supra* note 25, at 189.
31 *Id.* at 21-28.
32 Boyd & Dadayan, *supra* note 5, at 6.
33 *Id.* at 6-9.

pressure on tax collections: laws limiting taxes based on factors including income, inflation, and population growth; and balanced budget rules.[34] Balanced budget rules may require the budgets that governors submit to legislators or that legislatures adopt not to exceed actual revenues.[35] These types rules can incentivize legislators to err conservatively in forecasting revenues, as over-forecasting revenues may force legislators to take unpopular measures like cutting services or hiking taxes. Rose and Smith (2011) find a significant relationship between balanced budget rules and lawmakers' tendency to underestimate revenue, as did Krause, Lewis, and Douglas, who tested whether the state legislature was required to pass a balanced budget.[36] Other policies, like economic stabilization funds or surplus general funds, create a sense of fiscal "slack," leading estimators to worry less about maintaining an underforecast.[37]

A binding budget requirement can lead to increased incentive to over-forecast, as lawmakers become wary of a locked-in budget.[38] The fiscal calendar differs from state to state—some states use annual budgets while others use biennial; some states estimate revenue as often as six or more times a year while others only release an official estimate once every two years; some states project out well beyond the current fiscal cycle while others do not.[39] In general, research finds that errors increase the farther ahead a forecast

---

34 Rose & Smith, *supra* note 14, at 190.

35 James M. Poterba, *State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Policies*, 102 J. OF POL. ECON. 799, 803 (1994).

36 Rose & Smith, *supra* note 14, at 193; Krause, Lewis, & Douglas, *supra* note 26, at 283.

37 Rose & Smith, *supra* note 14, at 189.

38 Krause, Lewis, & Douglas, *supra* note 26, at 283.

39 *Id.*

is issued from the start of the forecast cycle, a finding that negatively impacts states using biennial budgets, since forecasters must determine revenue estimates for the second year over a year in advance.[40] More frequent forecasts do not necessarily yield more accurate estimates.[41]

4. Institutional Factors: Consensus Forecasting

Consensus-based forecasting, in which representatives of the executive and legislative branches of government join to agree on an estimate, is at the center of a long normative debate. Scholars differ on the value of consensus groups. A Pew-Rockefeller Institute study found no clear evidence that consensus-based processes produced more accurate estimates.[42] However, others extol the virtues of consensus-based programming, with a report by the Center on Budget and Policy Priorities describing consensus forecasting as a "common-sense best practice."[43]

Consensus forecasting may not add substantially to accuracy, but proponents claim that consensus-based methods produce greater political buy-in, making legislators more likely to abide by estimates produced by the process.[44] Proponents also claim that

40 Boyd & Dadayan, *supra* note 5, at 27, 33.
41 *Id.*
42 THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *supra* note 19, at 5.
43 McNichol, *supra* note 3, at 1.
44 *Id*. at 4; *see also* Mikesell & Ross, *supra* note 25, at 191.

consensus forecasting is correlated with an increase a state's credit rating, lowering its interest rates.[45]

In terms of optimal design for a consensus group, Krause and Douglas (2013) assert that consensus groups have a trade-off between levels of diversity and group size. They claim, in other words, that as a group increases in size and increases in diversity, its output quality decreases, but if the size increases but diversity decreases a more optimal bargain is struck.[46]

---

45 Jeffrey M. Tebbs, *Breaking the Stalemate: A Proposal for a Consensus Revenue Forecasting Process,* CONNECTICUT VOICES FOR CHILDREN 1 http://www.ctvoices.org/sites/default/files/bud09revenueforecasting.pdf (March 2009).

46 Krause & Douglas, *supra* note 23, at 139.

II. Group Decisionmaking in Behavioral Economics Literature

Behavioral economists have studied group decisionmaking in order to determine when and what types of groups make accurate or successful decisions. Much of this study has focused on one type of group: the deliberative group. Deliberative groups are here defined as groups that "[use] deliberation and [ask] for the reasoned exchange of facts, ideas, and opinions."[47] These groups stand in contrast to individual decisionmakers, or groups that do not do not permit consultation with other group members before coming to a decision (e.g., opinion polls).

Deliberative groups' output may be biased by a number of systematic fallacies. The way these fallacies affect groups' output depends on the nature of task the group is called on to complete. The first section will outline a taxonomy of deliberative group tasks, and argue that revenue estimating groups complete estimation tasks. The second section will describe how group fallacies identified by behavioral economists affect groups' accuracy in completing estimation tasks.

A. Taxonomy of Deliberative Group Tasks

Deliberative groups can perform multiple types of cognitive assessments using several response formats. Stasser and Dietz-Uhler developed a method for categorizing groups according to the nature of the response the group must produce (see Table 1).[48]

---

[47] Cass R. Sunstein. *Group Judgments: Statistical Means, Deliberation, and Information Markets*, 80 N.Y.U. L. REV. 962, 963 (2005).

[48] Garold Stasser & Beth Dietz-Uhler, *Collective Choice Judgment, and Problem Solving, in* BLACKWELL HANDBOOK OF GROUP PSYCHOLOGY: GROUP PROCESSES 34 (Michael A. Hogg and R. Scott Tindale, eds., 2001).

Selecting groups must chose two or more responses, while rating tasks involve determining a value along a continuum.[49] An example of a selective group is a contract management team assembling a list of acceptable vendors. An example of a rating task is a group of managers determining the best vendor to supply technology products from that list. These two categories are not wholly distinct, as part of selection entails rating individual options.

The second dimension of this schema is a continuum running between intellective tasks that correspond to an objectively correct answer or factual determination, and a judgmental task that has a subjective, not objective, answer.[50] Satisfying at least some of these conditions suggests the group's task is more intellective than judgmental: (1) group members share an inference system or procedural knowledge to determine the right answer; (2) the system of inference has enough information to determine a right answer; (3) individuals who know the right answer must be capable and willing to demonstrate how they arrive at that answer; and (4) those who do not know the right answer must accept experts' demonstration of correctness.[51]

Table 1. Group Categorization Schema.

	Response Format	
	Select	**Rate**
Judgmental	Choice	Judgment
Intellective	Problem Solving	Estimation

49 *Id.* at 32.
50 *Id.* at 33.
51 *Id.*

For instance, deliberative polling groups perform a judgmental task. The deliberative polling technique developed by Fishkin entails a representative jury of peers making a determination on a policy area, like whether to adopt a new currency or how to regulate utilities.[52] Because group members must make a subjective judgment based on their values and ideology, many possible "right" answers may exist, and members may not share the same system of inference.

Revenue estimation falls within the "Estimation" cell of the schema in Figure 1, as a rating group performing a mainly intellective task. These groups must rate, because they are responsible for returning a single revenue estimate figure from a continuum of possible revenues. They also must return a single, objectively right answer: the estimate, in theory, should correspond as closely as possible to the actual revenue taken in by the state. In contrast, elected officials are then responsible for the task of writing the budget, which is a value-judgment task more likely to fall in the "judgment" or "choice" cells. State budgets are political documents, with potentially multiple right answers depending on how lawmakers weight competing priorities.

Deliberative groups exhibit a special susceptibility to certain biases. As a subset of deliberative groups, behavioral psychologists have developed theoretical models and empirical models specific to groups performing estimation tasks. Intellective tasks like estimation permit scholars to explore the quality of group decisionmaking empirically by measuring group factual determination against the objectively correct outcome. This

---

[52] Robert C. Luskin, James S. Fishkin, & Roger Jowell, *Considered Opinions: Deliberative Polling in Britain*, 32 BRITISH J. OF POL. SCI. 455, 461-62 (2002).

section will focus on the systematic biases exhibited by deliberative groups, particularly with reference to estimation tasks.

B. Deliberative Groups Biases

Deliberative groups exhibit systematic patterns of error biasing the accuracy of their factual determinations. Sunstein (2015) describes two forces generating these errors: group members misconstruing the informational signals sent by others, and social pressures to keep quiet about contrary viewpoints.[53] These two forces in turn create four general types of problems: (1) cascade effects, in which groups blindly follow the lead of those speaking first; (2) amplification of the errors made by individual group members; (3) group polarization in favor of a more extreme version of a belief already previously held by group members; and (4) overemphasis on the group's common knowledge at the expense of hidden (but crucial) knowledge.[54]

Generally, each of these errors relates back to the group's inability to draw out the most important information, as well as shortcomings in the group's analysis of its limited information. These mechanisms explain why deliberative groups come to wrong factual determinations – a crucial challenge for groups attempting to project accurate revenue forecasts. This section will define each error with greater detail, and identify observable variables affecting a group's propensity to fall prey to each of these errors.

---

53 CASS R. SUNSTEIN & REID HASTIE, WISER: GETTING BEYOND GROUPTHINK TO MAKE GROUPS SMARTER 22-23 (2015).
54 *Id.* at 23-24.

Cascade effects, the first type of group error, occur when individuals act against their internal, independent judgments to side with the first opinions they observe others express.[55] Various types of cascades exist. In informational cascades, group members fall in line with the first opinion expressed because they assume the first member to speak has superior information.[56] In these cascades, every member fails to disclose their internal independent judgment, as well as any private information they hold.[57] In reputational cascades, a similar effect occurs because other group members do not want to lose face by disagreeing with the first opinion.[58]

Several variables may affect the propensity of groups to experience cascades, including the size of the group. One study attributed the reason smaller juries have the same mistrial rates as larger juries to information cascades. Because information cascades exert stronger effects in larger groups, argue Luppi and Parisi, larger juries actually had inflated rates of unanimity.[59] Cascades may also rely on the deliberation methods used by the group – do members formulate and submit their independent judgments for the entire group to evaluate? In addition, decision rules affect the existence of cascades – compared

---

55 Sunstein, *supra* note 42, at 999-1004.
56 *Id.* at 999-1000.
57 *Id.*
58 *Id.* at 1002.
59 Barbara Luppi & Francesco Parisi, *Jury Size and the Hung-Jury Paradox*, 42 JOURNAL OF LEGAL STUDIES 2, 399, 409 (June 2013).

to consensus or super-majority requirements, group that operate by majority rule exhibit almost no cascade effects.[60]

Second, a group setting can amplify individuals' patterns of misjudgment. To take one example, individuals show susceptibility to the representativeness heuristic, in which the apparent resemblance or similarity of two events will affect assessments of probability.[61] Individuals exhibit many other types of heuristics as well. While individuals will make certain predictable errors of judgment, groups will often operate to amplify at least *some* types of errors.

Some studies suggest that the extent to which a group amplifies individual errors depends on the mix of members in the group. Groups that amplify individual errors usually do so under the influence of logically or statistically irrelevant information, or because relevant information lacks influence.[62] Group members' individual susceptibility biases may predict the group's susceptibility as a whole:

> "[W]hen a bias arises due to a widely shared judgmental heuristic or belief system, group interaction will enhance the bias. In contrast, when the underlying cognitive process is less widely shared and groups are likely to contain one or more members who are not susceptible to the

60 Reid Hastie & Tatsuya Kameda, *The Robust Beauty of Majority Rules in Group Decisions,* 112 PSYCHOLOGICAL REVIEW 494-508, 495 (2005).

61 Sunstein, *supra* note 42, at 990.

62 Stasser & Dietz-Uhler, *supra* note 43, at 49.

> bias, the group interaction may provide an opportunity for more accurate members to persuade (or correct) less accurate members."[63]

More cognitively heterogenous groups have a greater chance of canceling out individual errors – at least, if those errors are not correlated. Individual errors with uncorrelated error terms include heuristics like the egocentric bias, in which individual decisionmakers believe that others think and reason as they do.[64] Individual errors with correlated errors present a much greater danger to groups, contributing to excessive group bias. Biases exaggerated by group dynamics include the representativeness heuristic, framing effects, overconfidence, and the sunk-cost fallacy.[65]

The third type of error is group polarization, a phenomenon connected to group cascades, which occur when members of a deliberative group end the deliberative process holding a more extreme version of the preconceived notions they initially held.[66] This phenomenon is demonstrated most clearly in judgmental tasks, in which a group must return a subjective determination, e.g., panels of judges or juries returning a verdict.[67] Behaviorists propose three main mechanisms to explain why groups polarize: (1) the initial position will, by nature of being first, have more arguments in its favor, lending numerical weight to the first opinions adopted; (2) members bow towards the initial opinions aired by other members out of deference to social pressures and perceived early

---

63 *Id.*
64 SUNSTEIN & HASTIE, *supra* note 48, at 54.
65 *Id.* at 52-53.
66 *Id*. at 77.
67 Sunstein, *supra* note 42, at 1004-05.

consensus; and (3) the most extreme members will have greater confidence in their views, and therefore speak more persuasively.[68]

The group polarization concept is harder to apply in an estimation task than a judgment task. To begin with, group polarization does not predict whether the group will return an accurate answer more or less frequently, since the group's final decision depends on the initial positions adopted by its members – this hinders the empirical testing of this phenomenon. In addition, in a setting where members' greatest concerns are accuracy instead of ideology, no connection may exist between members' extreme positions and their confidence in their conclusions.

Persuasive arguments theory suggests how this mechanism might work in an estimative setting. This theory predicts that, because individuals do not have access to full information, that "polarization occurs in the direction of the alternative for which the greatest number of arguments have been made."[69] Numerically counting the arguments in favor and against a lower or higher rating could predict the outcome adopted by the group – however, this requires a detailed level of access to the group's deliberative process.

Fourth, hidden profiles and common knowledge may adversely impact the quality of a group's output, failing share information as efficiently as possible to come to the most accurate conclusion possible. Groups do not aggregate information correctly when they disproportionately weight knowledge shared by all group members, a bias known as

---

68 *Id.*

69 Glen Whyte & James K. Sebenius, *The Effect of Multiple Anchors on Anchoring in Individual and Group Judgment*, 69 ORG. BEH. AND HUMAN DECISION PROCESSES 75, 78 (January 1997)

the “common knowledge effect.”[70] These groups do not correctly take into account the accurate information privately held by a few individuals, “hidden profiles” obscured by the common knowledge effect.[71] Studies suggest that the larger the group size, the more likely the group will exhibit the common knowledge effect by discussing shared information more and non-shared information less.[72]

Observers of this phenomenon suggests two explanations. First, similar to group polarization, the group may discuss common knowledge more frequently, numerically speaking, than hidden profile information, giving it greater weight.[73] Second, shared knowledge affects individuals’ judgments, which in turn affect the group’s judgment.[74]

C. Deliberative Group Strengths and Strategies for Improvement

The previous section delineated a long list of blind spots exhibited by deliberative groups, particularly groups performing estimation tasks. Despite their record for error, deliberative groups do have certain desirable attributes. Deliberative groups also demonstrate unique strengths, and may be amenable to changes significantly improving the quality of group decisions.

One strength of a deliberative groups is in conferring democratic legitimacy on an outcome – in other words, the act of deliberation can form a democratic end in itself, rather than as a means to that end. Dryzek (2001) conceptualizes legitimate democratic

---

70 Sunstein, *supra* note 42, at 994.
71 *Id.*
72 Garold Stasser *et al.*, *Information Sampling in Structured and Unstructured Discussions of Three- and Six-Person Groups,* 57 J. PERSONALITY & SOC. PSYCH. 67, 73 (1989).
73 Sunstein, *supra* note 42, at 998.
74 *Id.*

outcomes as decisions that are "resonant" with one or several competing, dispersed, broad-based discourses in a deliberative society.[75] Deliberative processes help group members rhetorically justify a decision to other members of society, by drawing on shared discursive frameworks to explain and rationalize decisions.

This concept is most clearly illustrated by groups performing judgmental (as opposed to intellective[76]) tasks, which require an application of shared values to determine a subjectively correct answer. One such group is a jury in a criminal or civil trial, in which the jury's verdict is lent legitimacy partly from the deliberative, justificatory method used to determine its decision. Another such group is Fishkin's deliberative polling method, in which a small, representative group of citizens determine through deliberation an ideal, binding policy solution (e.g., the desirability of imprisonment to combat crime).[77]

Even for groups performing estimative task requiring accuracy, deliberative group decisionmaking processes can generate a valuable legitimacy. Mikesell and Ross (2014) illustrate how this legitimacy operates in a state revenue forecasting context. They note that a consensus-based revenue estimating process incorporating several relevant political players in Indiana generated what they describe as a vital "acceptance" of budget estimates:

---

75 John S. Dryzek, *Legitimacy and Economy in Deliberative Democracy*, 29 POL. THEORY 651, 661-62 (2001).

76 *See* Table 1.

77 Sunstein, *supra* note 42, at 1009.

> The process installed … by agreement and not by statute … seeks to produce an accepted forecast that neutralizes political influences (both real and accused) on the forecast by involving all branches and both political parties. […] [A]cceptance of the fiscal baseline is a test of the forecast *process*, just as accuracy is critical for the fiscal baseline. In the years since the development of the Indiana consensus revenue forecasting approach, there have never been dueling forecasts produced within the General Assembly. *The Indiana legislature and governor have always accepted the forecast coming from the [revenue estimating body] as the one used for determining appropriations.* As a result, it served as a hard budget constraint.[78]

The deliberative and broad-based forecasting process induces acceptance of a budget recognized as legitimate, in the same way a verdict produced by a representative, deliberative jury has a broad legitimacy in society.

Legitimacy can salvage the value of deliberative groups. However, as described in part IIB, these groups still experience systematic errors of reasoning, distorting the accuracy of their outcomes. Behavioral scholars have proposed some solutions ameliorating the worst of these errors. Many of the heuristics worsening the quality of deliberative group outcomes are rooted in two sources: misconstrued informational signals and social pressures in favor of conformity.

---

78 Mikesell & Ross, *supra* note 25, at 200 (italics original).

Various group techniques ameliorate these sources of bias. For instance, the "nominal group" technique encourages members of the group to submit anonymous recommendations and then serially discuss these recommendations.[79] Anonymity limits social pressures to conform behind the best-regarded expert. Multiple rounds generate multiple opportunities to disclose information, encouraging group members to use the best information available and discouraging hidden profiles.

A similar method of making decisions is the Delphic judgment technique, in which anonymous panels of experts contribute their best estimates over several rounds, eventually converging on a group consensus around the best answer.[80] Deliberation between group members is typically highly structured, with members' individual conclusions potentially mediated by a group facilitator, and the final judgment is a statistical aggregation of members' individual judgments.[81] Because this method incorporates less deliberation, these groups may experience a lessening of the social pressures leading to group polarization – the most vocal and confident members are muted by the mediation of feedback.

Altering group membership to include more or stronger voices from people with access to the correct answer tends to increase group accuracy. When a person confidently

---

79 For a definition and example of nominal group techniques, *see* P. Delp *et al.*, *Nominal Group Technique*, SYSTEM TOOLS FOR PROJECT PLANNING, International Development Institute, http://www.aucd.net/docs/urc/Leadership_Institute/Subsequent%20Leadership%20Institute%20Materials/Nominal%20Group%20Technique.pdf.

80 Gene Rowe & George Wright, *Expert Opinions in Forecasting: The Role of the Delphi Technique*, *in* PRINCIPLES OF FORECASTING, 125, 125 (J.S. Armstrong, ed., 2001).

81 Sunstein, *supra* note 42, at 1018.

relates a self-affirming or clearly correct answer (e.g., estimating distance between London and Paris), the group will likely accept that answer.[82] Social pressures likely induce conformity in favor of the most confident person, who seems to act on expertise or prior knowledge.

D. Alternatives to Deliberative Groups

The second set of solutions concerns jettisoning the deliberative group structure altogether in favor of alternative decisionmaking structures promoted by behavioral scientists. One of these alternatives is the "statistical group," or polling a random selection of people and taking the mean answer. Another alternative is the "information market," in which many members contribute individually held information to a final product (e.g., a Wikipedia article or a market price).[83]

Statistical groups consistently show greater accuracy than deliberative groups,[84] in part because individual propensity for error is canceled out rather than exaggerated through deliberation. In a meta-analysis of over a dozen studies testing the factual accuracy of group decisionmaking, Gigone and Hastie (1997) concluded that deliberative groups rendered decisions approximately as accurate as the mean of all members polled: "For the most part, group judgments tend to be more accurate than the judgments of typical individuals, approximately equal in accuracy to the mean judgments of their

---

82 *Id.* at 1007.
83 *Id.* at 1022-23.
84 *See, e.g., id.* at 971.

members, and less accurate than the judgments of their most accurate member."[85] This finding makes some sense: statistical groups combine each individual's best information, without the distorting social pressures of deliberation.

Information markets work similarly, in that individuals aggregate information without the distorting social pressures of deliberation. In fact, some information markets may encourage an even higher quality of information – an investor who acts on her guess at the true value of a stock by buying or selling likely has a higher degree of confidence and more motivation to seek information than a random polled respondent.

---

85 Daniel Gigone & Reid Hastie, *Proper Analysis of the Accuracy of Group Judgments,* 121 PSYCH. BULLETIN 149, 153 (1997).

III. Applying Behavioral Economics insights to State Revenue Estimation

Insights from behavioral economics may help explain which state revenue estimation processes perform best in terms of accuracy, particularly for states relying on deliberative groups to determine a revenue estimate. In revenue estimation, the consensus-based process fits most neatly within the category of an estimative deliberative group: members must perform a rating task, agreeing on a single revenue estimate to begin the budgeting cycle.

A. Generating Hypotheses

1. Comparing Use of Consensus Groups to Other Methods

According to behaviorists' account, deliberative groups magnify individuals' biases to generate less accurate answers as a collective.[86] Deliberative groups also fare worse than other methods of agglomerating individuals' judgments, including statistical sampling and information markets. Many states incorporate a strong deliberative aspect to consensual revenue estimation methods, as illustrated by this description of Indiana's process:

> [T]he forecasting group does not merely bring together competing forecasts developed by the representatives on the committee but also involves the forecasting group in developing the methodology employed to create the forecast. In contrast, for instance, the Michigan process produces consensus

86 *See* Section IIB on amplification of individual biases.

from alternative economic and revenue forecasts…, with each independently developed [by a separate agency].[87]

The Indiana Revenue Forecast Technical Committee resembles a deliberative group, as its members do spend time in discourse with each other before submitting their final recommendation. This time spent in discussion is an open opportunity for group members to become more polarized or amplify others' biases.

Hypothesis 1: Consensus-based deliberative groups perform less accurately than non-consensus based groups, including states in which executive branches or separate branches forms of estimation.

2. *Comparing States Within the Consensus Group Category*

Section IIB identified several relevant variables that could affect the performance of a deliberative group detrimentally, whether by increasing social pressures or by suppressing high-quality information. As described in that section, deliberative groups are more likely to experience an informational cascade in larger groups and when groups require unanimity rather than majority agreement. Individual biases are less likely to be amplified when groups are more cognitively heterogeneous; similarly, groups are more likely to over-weight common knowledge when a large number of group members come from a similar background.

If these insights are applicable to the revenue estimating context, then the size of forecasting groups, the makeup of the groups, and the decisional rules they operate under

[87] Mikesell & Ross, *supra* note 25, at 195.

matter. This theory was explored by Krause and Douglas (2013). They hypothesized that the forecasting consensus group would improve in accuracy as the group's diversity and size became increasingly negatively correlated.[88] In other words, groups that were large and extremely diverse would not perform well; nor would small and homogenous groups. Only when groups had mixed characteristics—large but homogenous, or small but diverse—would the group's performance improve. They argued that, as groups increased in size and heterogeneity, the variety eventually imposes collaboration costs on group members that are too high. Krause and Douglas built and tested a regression analysis that controlled for four different types of diversity, the size of the group, the decisional rule the group used (e.g., unanimity, majority, etc.) and so on. The decisional rules did not test as statistically significant, but the variables testing the interaction between various measures of group diversity and group size did.[89]

> *Hypothesis 2:* Revenue estimation groups operating under majoritarian rules will produce more accurate revenue estimates than groups with unanimity voting rules.

Deliberative groups exacerbate patterns of bias by increased social pressure. Strict decisional rules, including unanimity requirements, may increase social pressure on members of the group who would otherwise raise important issues, which in turn may increase the chances of bias amplification through pressures like group polarization or cascade effects. While Krause and Douglas did not find statistically significant effects in

88 Krause & Douglas, *supra* note 23, at 139.
89 *Id.* at 143.

their regression models, this descriptive analysis will be more exploratory in nature, and will not claim to find a causal relationship.

> *Hypothesis 3:* The relative accuracy of revenue estimation will depend on group size, with larger groups more likely to produce larger error terms than smaller groups.

As outlined in Section IIB, larger deliberative groups are more susceptible to experiencing information cascades, in which the group inefficiently uses the information it has at hand. This test differs slightly from the Krause and Douglas test, who mainly sought to test the interaction between group size and various measures of diversity.

B. Data and Methods

Previous analyses have shown that state-level revenue estimators exhibit a well-established conservative preference to underestimate the amount of tax revenue in a given budget cycle, generally for political reasons.[90] This presents some difficulty: how can forecast error attributable to political pressures be separated from forecast error attributable to forecasters' collective mistake?

One way to do so might be to examine how revenue estimators forecast in unusually difficult circumstances. In particular, the period immediately preceding the onset of a recession presents great difficulty for forecasters, as evidenced by substantial overestimation errors during the last two recessions. Projecting just enough—or too little—in tax revenues "might be a sign of fiscal stress because these estimated revenues

90 Williams & Calabrese, *supra* note 9, at 130.

are needed to cover immediate spending... These errors may not reflect political decisions."[91] In normal years, forecasters' error terms are small and positively biased, indicating a preference for underestimating.[92] But in recent recessions, state forecasters not only produce negative forecast errors, the absolute value of those errors has substantially increased as well. In addition, more states misestimate recessions—the proportion of states overestimating revenue increased from 25 percent in the 1990-92 recession to 45 percent in 2001-2003 to over 70 percent by 2009.[93]

Ideally, revenue forecasters in deliberative groups should outperform mechanistic economic models because of their capability to judge the future and their access to high-quality information about economic trends. For instance, in explaining their measure of forecast difficulty, Boyd and Dadayan (2014) note that, in contrast to their naïve model drawing only on available economic data and past trends, human forecasters should have an advantage in determining turning points and changes in the economy's direction: "Because revenue forecasters read the newspapers, talk to economic forecasters, and make use of a wide variety of information sources that cannot easily fit into uniform models, they are likely to be more accurate than our naïve model."[94]

For these reasons, the best way to assess the quality of state revenue forecasters' work is to focus on the size of their forecast errors during recessionary years. Because

---

91 *Id.* at 131.

92 Boyd & Dadayan, *supra* note 5, at 15.

93 THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *supra* note 52, at 8.

94 Boyd & Dadayan, *supra* note 5, at 14.

state forecasters generally attempt to underestimate tax revenues for political reasons, years in which they produce negative errors represent a deviation from the norm. High-performing revenue estimating groups should have the capability to gather high-quality information from a variety of sources, as described by Boyd and Dadayan, to anticipate a possible downturn. State forecasters with the lowest error rates during those years demonstrate an unusually high level of accuracy.

In addition, state forecasters have particular difficulty in estimating certain especially volatile revenue streams. Three sources of revenue make up 72 percent of the average state's total revenue: personal income tax, sales tax, and corporate income taxes. Of these three, the corporate income tax consistently demonstrates the greatest volatility. State forecasters mis-estimated corporate income tax revenues by an average of 1.3 percent with a standard deviation of 22.6 percent, almost three times as much variation as for the personal income tax and 5 times as much for the sales tax.[95] Individual and collective discernment and judgment is key for measuring volatile revenues in uncertain circumstances.

Dependent variables. States self-report forecasts and revenues through the twice-annual *Fiscal Survey of the States* report, collected by NASBO and the National Governor's Association. This analysis will draw on reported projections and collections

---

95 *Id.* at 7.

of the corporate income tax from the Fall 2007 report to the Spring 2011 report, to coincide with the beginning and end of the most recent recession.[96]

A frequently used method to measure forecasting error is mean absolute percentage error (MAPE), where F_t refers to the forecasted projection and A_t refers to the actual revenue projection in time t[97]:

$$\text{MAPE} = \sum_{t=1}^{T} \frac{|\mathrm{F}_t - \mathrm{A}_t|}{\mathrm{A}_t}$$

MAPE measures the magnitude of the difference between the forecasted and actual revenues, but does not show bias as clearly. A better measure for understanding the direction in which forecasters err is the mean percentage error (MPE)[98]:

$$\text{MPE} = \sum_{t=1}^{T} \frac{\mathrm{F}_t - \mathrm{A}_t}{\mathrm{A}_t}$$

The MPE becomes negative when revenues are less than forecast and positive when revenue exceeds forecasts. Figure 1 demonstrates this graphically: at the height of the

---

96 NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS & THE NATIONAL GOVERNORS ASSOCIATION, *The Fiscal Survey of States* (Fall 2008-Spring 2011) https://www.nasbo.org/mainsite/reports-data/fiscal-survey-of-states/fiscal-survey-archives.
97 Mikesell & Ross, *supra* note 25, at 196; *see also* Boyd & Dadayan, *supra* note 5, at 6.
98 Mikesell & Ross, *supra* note 25, at 196.

recession in 2009, forecasters issued projections 5-25 percent above actual revenues for each of the three major tax streams. The chart also illustrates the varying levels of unpredictability for each of the three funding sources, with sales tax forecasts tracking more closely with actual revenues than corporate income tax forecasts.

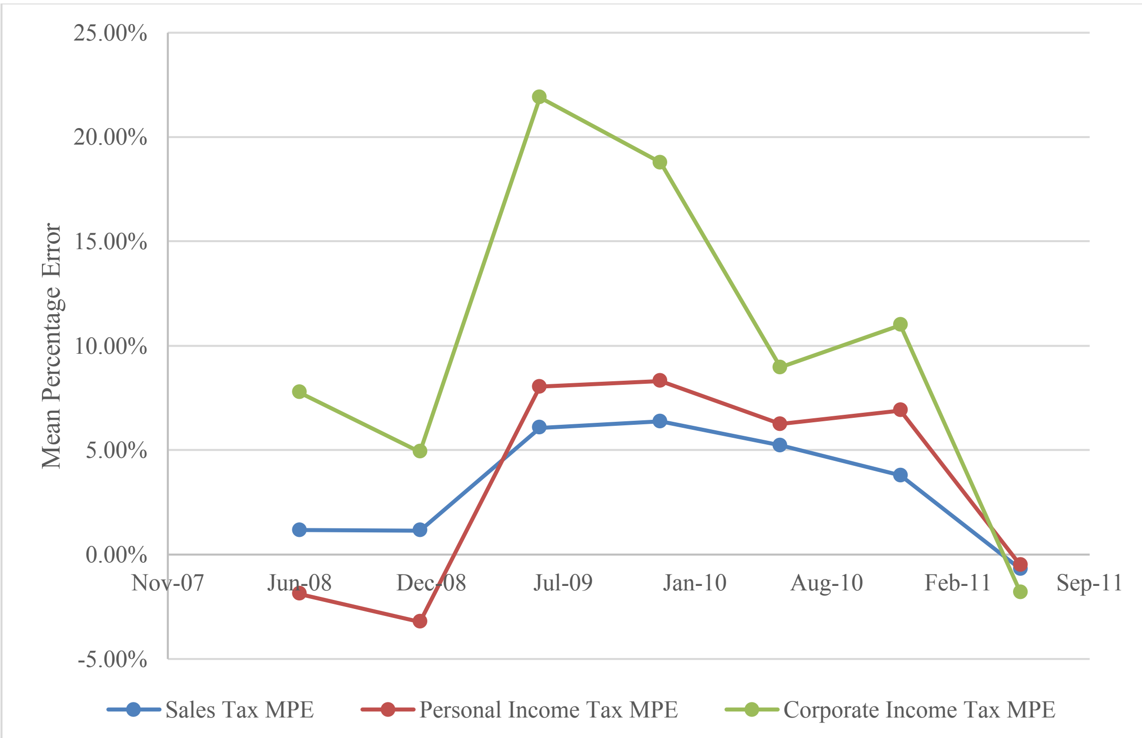

Figure 1. Mean Percentage Error for Tax Revenue Streams, 2008-2011

Independent Variables: Hypothesis 1. Hypothesis 1 requires a comparison between the relative performance of states using consensus groups for revenue estimation, states in which separate branches generate estimates, and states where the executive branch alone issues projections. Of the 45 states that collect corporate income tax, 18 use consensus estimation, 5 use separate estimation, and 10 use executive branch estimation. The remaining 12 states were not unambiguously classifiable as one of those three categories.[99]

Independent Variables: Hypothesis 2. Krause and Douglas (2013) gathered decisional rules for twenty-six states operating under the consensual group system. Each state was categorized as either a “unanimity” state or a “majority” state.[100]

Independent Variables: Hypothesis 3. Krause and Douglas (2013) also collected information on the panel size used by twenty-six consensus group states. The size of panels ranged from 3 to 9, with Delaware as an outlier with a 30-person consensus group panel.[101]

Method. The MAPE and MPE measures provide a rudimentary way of evaluating the accuracy of forecasters’ projections. However, forecasters’ true performance remains obscured by noise. A common method of measuring forecast difficulty is to compare forecasters’ performance against naïve forecast models using simple decision rules.[102] Following the practice of Mikesell and Ross (2014),[103] these two simple models provide a yardstick against which to measure state forecasters’ performance:

1. A basic lag model setting the state’s forecast F_t at time *t* equal to the state’s actual revenue A_{t-1} in time *t-1*:

$$\mathrm{F}_t = \mathrm{A}_{t-1}$$

99 *See* Appendix A.
100 Krause & Douglas, *supra* note 23, at 140.
101 *Id.*
102 Boyd & Dadayan, *supra* note 5, at 12-15; Mikesell & Ross, *supra* note 25, at 196-97.
103 *Id.*

2. A simple trend model setting the state's forecast F_t at time t equal to a univariate regression of observed values:

$$F_t = \beta_0 + \beta_1 t$$

The next section will set out the states' average MAPE and MPE forecast errors in calculating the corporate income tax, measure the forecast errors against the two simple forecast models, and break out certain sub-groups of interest.

C. Analysis

Between 2008 and 2011, state governments experienced an extraordinary amount of revenue volatility. Unlike other recessions, all three major sources of revenue—sales tax, personal income tax, and corporate income tax—declined simultaneously. Table 3[104] shows the extent of the downturn; none of the MPE descriptive statistics carries a positive sign, indicating that forecasters profoundly overpredicted revenues. Some states did not fare so poorly. North Dakota, for instance, averaged strong negative MPEs in all tax revenue categories for this period, with the lowest average annual sales tax MPE in the nation at -6.94 percent. The example of

Table 2. Descriptive Statistics of Tax Revenue Streams, 2007-2011.

	Average Sales MPE	Average Personal Income Tax MPE	Average Corporate Income Tax MPE
Mean MPE	3.30%	3.42%	10.09%
Standard Deviation	3.24%	4.50%	16.51%
75th Percentile	4.39%	6.08%	16.23%
Median MPE	3.70%	3.89%	6.34%
25th Percentile	2.30%	2.05%	1.73%
n	318	291	313

104 Author calculations; *see* Appendix B; NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS & THE NATIONAL GOVERNORS ASSOCIATION, *supra* note 90.

North Dakota illustrates the value of comparing state forecaster error with error from naïve models—the negative error rates do not reflect state forecasters' perspicacity so much as an exogenous, unforeseen economic boom precipitated by technological advancements in energy extraction.[105]

Focusing in on the revenue stream of interest, the corporate income tax, demonstrates some of the difficulty of forecasting during this time. Not only was the magnitude of corporate income tax errors large, but measures of variance also indicate a very wide range of error values. Half of states' average corporate income tax errors during this period ranged between 1.73 and 16.23 percent, a very wide spread indicating estimators' relative unpreparedness for declining tax receipts.

Finally, Table 4[106] demonstrates how comparing forecasters' projections with naïve models helps isolate forecasters' value in terms of accuracy. Columns (1) and (2) show two types of forecaster errors, MPE and MAPE. MPE can carry a positive sign, indicating that forecasters' projections overestimated tax revenue, or a negative sign, indicating that estimators issued an underestimate of tax revenues, as reflected in North Dakota. MAPE, in contrast, is an absolute

105 Jack Healey, *Built Up By Oil Boom, North Dakota Now Has an Emptier Feeling*, N.Y. TIMES (Feb. 7, 2016) https://www.nytimes.com/2016/02/08/us/built-up-by-oil-boom-north-dakota-now-has-an-emptier-feeling.html.

106 Author calculations; *see* Appendix B; NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS & THE NATIONAL GOVERNORS ASSOCIATION, *supra* note 90.

Table 3. Mean Absolute Percentage Error (MAPE), Mean Percentage Error (MPE), Naïve Models (1) and (2), and Forecaster comparisons (3) and (4) for state corporate income tax receipts, 2007-2001.

	(1) MPE	**(2) MAPE**	**(3) Naïve Lag MAPE**	**(4) Naïve Trend MAPE**	**(5) Naïve Lag MAPE - State MAPE**	**(6) Naïve Trend MAPE - State MAPE**
North Dakota	-15.36%	30.86%	23.29%	27.51%	-7.57%	-3.35%
National Average	4.58%	18.97%	17.08%	21.71%	-1.89%	2.74%

value measure reflecting the magnitude but not the bias of the error. Columns (3) and (4) show the forecast errors of the two naïve models. The naïve lag model uses actual revenue at time t_{i-1} as a forecast for actual revenue at time t_i, while the naïve trend model reflects the difference between actual revenue and the results of a simple univariate regression fitted on past revenue values.

Finally, columns (5) and (6) show the difference between the forecasters' actual MAPE error and the MAPE error generated by each of the naïve models. A positive value in those cells indicates that the state forecasters' projections fit the actual revenue better than the naïve model; a negative value shows that the naïve model surpassed the forecasters' projections in accuracy; and a value near zero means that the state forecasters' projections and the naïve model came to very similar estimates. In the case of North Dakota, both cells in columns (5) and (6) are negative, indicating that the naïve models performed far better. The forecasters in North Dakota did not add additional value in terms of accuracy to what extant economic trends already indicated—the forecasters could not anticipate the additional revenues generated by the sudden energy boom. In contrast, the national average state forecast error tracked the two naïve models more closely, with the average state forecaster outperforming the naïve trend model.

1. Comparing Consensus and Non-Consensus States

Behavioral economics would tend to predict worse performance from deliberative groups in consensus estimation states, where representatives of multiple agencies or branches of government must meet and generally deliberate before adopting a starting figure for the budget. The literature would tend to predict that a single individual or entity would produce a more accurate output, unbiased by social pressures or inefficient use of information endemic to deliberative groups. At best, if the judgment of multiple people must be aggregated, Sunstein (2005) advocates instead for the superior accuracy of statistical groups or information markets, in which individuals submit their opinion with no opportunity to discuss with others.

The methods non-consensus states use to settle on a revenue estimate bear some similarity to statistical groups or information markets. For instance, in some states separate branches of government produce revenue estimates, with a governor and a legislature submitting divergent projections of total tax collections. New Jersey, for instance, has a system in which the branches separately come to conclusions about an appropriate revenue estimate. The Governor presents a budget estimate to the legislature, which then reviews and adjusts these estimates.[107] However, the legislature's power is strictly limited, as the Governor has the power to modify the appropriations bill through line-item veto and sole authority to certify the budget.[108] The process is contentious, with

---

107 Richard F. Keevey, *Budget Basics: How the State Budget is Developed and Revenue Estimated,* NJ SPOTLIGHT (Oct. 13, 2017) http://www.njspotlight.com/stories/17/10/12/budget-basics-how-the-state-budget-is-developed-and-revenue-estimated/.
108 *Id.*

occasional gaps between the two branches' determinations, and the rules do not especially incentivize a discursive rapport between the legislature and the governor.[109] However, returning to behavioral economics, increasing discourse could have the effect of biasing the independent determinations of the legislators and the governor. Hypothesis 1 posits that non-consensus states, especially in the executive and separate branch systems, may perform better and more accurately than consensus-based states in revenue estimation.

Table 5 sets up a comparison between the forecasting performance of consensus states and non-consensus states. Consensus state forecasters did make larger errors in estimating corporate income tax revenue than non-consensus forecasters. Both types of forecasters overestimated incoming revenue during this time, as indicated by the positive sign on the MPE estimates, but consensus forecasters did so by a larger margin per the MPE and MAPE averages. In addition, while both types of forecasters underperformed

Table 4. Comparison of Consensus States and Non-Consensus States in Estimating the Corporate Income Tax, 2007-2011.

	n (states)	**(1) MPE Average**	**(2) MAPE Average**	**(3) Naïve Lag MAPE Average**	**(4) Naïve Trend MAPE Average**	**(5) Naïve Lag MAPE - State MAPE**	**(6) Naïve Trend MAPE - State MAPE**
Consensus State	18	6.34%	19.24%	15.66%	22.37%	-3.58%	3.13%
Non-Consensus	16	3.17%	15.18%	14.52%	16.46%	-0.66%	1.28%

109 *Id.*

relative to the naïve lag model, non-consensus states performed much closer to the model's estimate, although consensus state forecasters achieved more accuracy than the naïve trend model. On the whole, though, this table tends to support the assertion of Hypothesis 1 that non-consensus groups will tend to outperform consensus-based groups.

For a more granular picture, Table 6 disaggregates the non-consensus category into executive branch and separate branch subcategories. Of the three types, executive branch states performed best, with consistently small error estimates across the table. The executive branch MPE estimate carries a negative sign, meaning that this category of states alone was successful in underestimating, rather than overestimating, revenue. Consensus states once again err by a greater margin than both other categories, and once again underperform relative to the naïve lag model by a wider margin than either executive or separate branch states.

Table 5. Comparison of Consensus States, Executive Branch States, and Separate Branch States in Estimating the Corporate Income Tax, 2007-2011.

	n (states)	**(1) MPE Average**	**(2) MAPE Average**	**(3) Naïve Lag MAPE Average**	**(4) Naïve Trend MAPE Average**	**(5) Naïve Lag MAPE - State MAPE**	**(6) Naïve Trend MAPE - State MAPE**
Consensus State	18	6.34%	19.24%	15.66%	22.37%	-3.58%	3.13%
Executive Branch State	5	-0.83%	15.57%	14.74%	16.11%	-0.83%	0.54%
Separate Branch State	11	4.99%	15.00%	14.42%	16.62%	-0.58%	1.62%

In some ways, though these findings are preliminary and require significance testing, the shortcomings of the consensus model are not surprising. Boyd and Dadayan (2014) note in their review of the literature and investigation of consensus process error that consensus forecasts do not yield particular accuracy. Even if consensus processes do not promise additional accuracy, they may yet be more desirable than executive and separate branch systems for other reasons. Separate branch systems often cause citizens frustration over the time spent debating revenue instead of policy matters. [110] Moreover, advocates claim a correlation between consensus forecasting and higher credit ratings, suggesting that ratings agencies give better ratings to consensus forecasting states.[111]

2. Comparing Consensus Group States

Decisional rules in a group can change the group's dynamic and the willingness of minority party members to speak up, to everyone's benefit or detriment. Hypothesis 2

Table 6. Comparing Unanimous and Majority Rule Consensus States' Accuracy in Estimating the Corporate Income Tax, 2007-2011

	(1) MPE Average	**(2) MAPE Average**	**(3) Naïve Lag MAPE Average**	**(4) Naïve Trend MAPE Average**	**(5) Naïve Lag MAPE - State MAPE**	**(6) Naïve Trend MAPE - State MAPE**
Unanimous	8.67%	21.61%	16.34%	24.01%	-5.27%	2.40%
Majority	4.00%	22.48%	18.99%	23.65%	-3.49%	1.17%

110 *Id.*, *see also* McNichol, *supra* note 3, at 5 ("in 2009, when Connecticut had a Republican governor and Democratic legislature, it took weeks—which could have been spent debating policy—to agree on a base revenue estimate"); Tebbs, *supra* note 40, at 1 ("As a result of this impasse, Connecticut's political leaders have wasted weeks wrangling over the size of the deficit rather than the difficult policy decisions necessary to bring the budget into balance").
111 *Id.* at 3.

argued that groups operating under a majoritarian system would produce more accurate predictions, as individual members would feel less constrained by a unanimity requirement and contribute better information.

In fact, Table 7 shows there is only very limited difference between the unanimous and majority error terms. Both the MPEs are positive, meaning the forecaster projected more revenue than actual collections. Majoritarian groups have a slightly larger absolute error, but the error terms for unanimous and majoritarian groups track each other closely. Neither group outperforms the basic naïve lag model, but both groups perform better than the naïve trend estimates.

Finally, Hypothesis 3 posited that larger consensus groups would produce less accurate revenue estimates than smaller groups, as a result of distortions like information cascades which are more likely to be experienced by larger groups. Consensus forecasting groups do not vary in size *too* greatly. Groups range from three to nine panelists, with the exception of Delaware, which uses a 30-person panel.

Table 8 shows the average errors experienced by panels of varying sizes. The first two rows take the average of all consensus states, with the exception of the outlier Delaware for the second row. The next five rows show the average errors for states using at each size band, from 3 to 7 panelists. The final two rows split the consensus states in two categories: relatively smaller panels of 3 or 4 panelists, and relatively larger panels of 5 to 9 panelists. The third column lists the number of states per category.

The findings from Table 8 tentatively confirm that larger states tend to experience bigger errors. Accuracy generally decreased as panel size increased, with the least accurate results from states with 6 panelists (although that category only contains two states, Kentucky and New York). Comparing the relatively smaller states of 3 or 4 panelists to the larger states with 5 to 9 panelists again shows that the relatively smaller states outperform larger states, with a smaller MPE of 6.91 percent and MAPE of 19.08 percent compared to larger panels with an MPE and MAPE of 9.25 percent and 24.15 percent respectively. Smaller states outperformed both naïve models, while larger states only outperformed the lag model, falling behind on the naïve trend model.

Table 7. Comparing Consensus State Panel Sizes and Relative Accuracy in Estimating the Corporate Income Tax, 2007-2011

Group	**Average Panel Size**	***n***	**MPE Average**	**MAPE Average**	**Naïve Lag MAPE Average**	**Naïve Trend MAPE Average**	**Naïve Lag MAPE - State MAPE**	**Naïve Trend MAPE - State MAPE**
All States (including outliers)	5.87	23	7.05%	21.91%	17.26%	23.88%	-4.65%	1.97%
All States (except outliers)	4.77	22	7.97%	21.38%	16.70%	23.38%	-4.68%	2.00%
3 Panelists	3	5	1.67%	11.74%	12.60%	14.49%	0.86%	2.75%
4 Panelists	4	6	11.21%	24.36%	16.49%	21.94%	-7.86%	-2.41%
5 Panelists	5	3	-1.00%	20.72%	18.66%	21.59%	-2.06%	0.87%
6 Panelists	6	2	31.21%	37.87%	15.06%	52.80%	-22.81%	14.94%
7 Panelists	7	3	11.79%	28.61%	21.60%	30.59%	-7.01%	1.98%
3-4 Panelists	3.54	12	6.91%	19.08%	16.37%	18.88%	-2.71%	-0.20%
5-9 Panelists	6.25	10	9.25%	24.15%	17.11%	28.79%	-7.05%	4.63%

D. Discussion

The behavioral economics literature shows that estimative deliberative groups tend to produce biased, inaccurate outputs. Deliberative groups are more likely to experience distortions for a number of reasons: cascade effects, in which the information or opinions raised earliest carry more weight; amplification of individuals' existing biases; polarization in favor of the most extreme opinions or opinions with the highest number of arguments in their favor; and weighting common knowledge held by the group over information held by individual members. These problems lead deliberative groups to systematically produce outputs that are more biased than outputs produced by individuals, and also underperform relative to other aggregation methods like statistical groups or information markets.

Consensus forecasting states use a process most similar to the deliberative group method tested by the behavioral economics literature. Consensus forecasting states use a process in which representatives from multiple branches of government meet in deliberative groups to settle on a single revenue forecast. These states contrast with executive forecasting states, in which the executive branch alone produces the forecast (usually through the governor's office or a dedicated budget agency) and states in which separate branches of government produce the forecast. Within the consensus forecasting category, groups have different characteristics, differing in terms of size, makeup, decisional standard, and other factors.

1. *Overview of Findings*

The analysis examined differences in accuracy between consensus and non-consensus states as well as within the consensus state category, measuring estimators' performance against data for the most difficult revenue stream to estimate: corporate income taxes during the Great Recession. This section also compared estimators' performance during this time to two naïve models of estimate error. The average state forecaster did not estimate corporate income tax more accurately than a naïve model using a lagged variable, although they outperformed a naïve model predicting revenue according to a simple linear trend.

Analysis demonstrated some limited evidence exists for the propositions that consensus groups will underperform non-consensus groups and that certain types of consensus groups will outperform other types. One of the findings of this paper was that consensus groups do not, on average, produce estimates any more accurate than states that have the executive branch or separate branches of government estimate tax collections. In accordance with predictions generated by the behavioral economics literature, consensus groups requiring unanimity outperformed majoritarian groups and smaller consensus groups performed better than larger consensus groups.

2. Directions for Future Research

There are several further directions for research coming out of this paper. These findings, while interesting, require~~s~~ a more robust test of its statistical significance and of the magnitude of the difference. Econometric modeling would help determine whether the relationship between these characteristics and forecasting accuracy is causal rather than merely associational. Perhaps future researchers can run experiments involving state forecasters to test claims in the behavioral economics literature that statistical groups and information markets outperform deliberative groups. will begin to experiment with some of the alternatives to deliberative groups, such as information markets or statistical groups. As all streams of revenue continue to increase in volatility, states need a stable and accurate source of information about their financial future.

3. Guidance for Policymakers

Policymakers tasked with designing institutions responsible for estimating revenue can draw on some of the findings above to maximize chances of estimate accuracy. These recommendations are made on the assumption that findings based on data drawn from recessionary years focusing on a single revenue stream generalize to other revenue estimation contexts.

First, policymakers should consider tasking revenue estimation to the executive branch alone or sharing the responsibility between executive and legislative branches, rather than relying on consensus groups to forecast revenue. Consensus groups produce larger misestimates than both executive and separate branch states, and also underperform relative to a basic naïve lag model by a greater magnitude.

This recommendation has some caveats. Notably, accuracy is not the only factor used to determine the desirability of different types of forecasting models. Lawmakers may also wish to ensure that a given revenue forecast is politically accepted by all major players in the budgetary process. Mikesell and Ross show that consensus-based forecasting may generate broad-based political buy-in for the revenue estimate, making the final estimate less likely to be assailed by other branches of government.[112] Consensus-based forecasting may deliver additional benefits—for instance, advocates

---

[112] Mikesell & Ross, *supra* note 25, at 200.

claim that an association exists between higher credit ratings and consensus-based forecasts, although this association is correlational rather than causal.[113] Finally, citizens may prefer executive branch forecasting to separate branch forecasting, as the latter can cause citizens frustrations with the protracted public wrangling over the budget.[114]

Second, if policymakers choose to continue using consensus groups to estimate revenue, they can maximize accuracy by tweaking some of the features of those groups. Consensus group states experience the lowest error terms with 3 to 4 panelists; more specifically, groups with 3 panelists demonstrate the lowest percentage of error. Finally, for greatest accuracy policymakers should permit these groups should use a majoritarian decision rule rather than requiring unanimous agreement.

---

[113] Tebbs, *supra* note 40, at 3.

[114] McNichol, *supra* note 3, at 5 ("in 2009, when Connecticut had a Republican governor and Democratic legislature, it took weeks—which could have been spent debating policy—to agree on a base revenue estimate"); Tebbs, *supra* note 40, at 1 ("As a result of this impasse, Connecticut's political leaders have wasted weeks wrangling over the size of the deficit rather than the difficult policy decisions necessary to bring the budget into balance").

Conclusion

State budget writers rely on accurate revenue estimates from forecasters. Overestimating the amount of incoming revenue can lead to service cuts or tax increases in states with balanced budget requirements, while underestimating revenue may result in collecting more taxes from citizens than state governments need. This report demonstrates that behavioral economics insights showing that deliberative groups perform poorly on tasks requiring accuracy hold true for revenue estimation groups too. Consensus-based groups consisting of representatives from multiple branches of government performed with less accuracy than non-consensus-based forecasting models. Within the category of consensus-based groups, smaller groups and majoritarian groups outperformed bigger groups and unanimous groups respectively, confirming predictions generated by the behavioral economics literature. These insights can help inform policymakers responsible for designing and managing revenue forecasting institutions.

Appendix A: State Revenue Estimation Typology

Several sources have compiled various lists categorizing states' revenue estimation practices. The list below was drawn out of three sources: Krause and Douglas (2013), NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS (2015), and McNichol (2014). Where all 3 sources agree that a state uses a consensus process, that state is classified as "consensus." Where McNichol (2014) classifies a state as executive branch or separate branches and no other sources claim the state belongs in a different category, that state is assigned to executive or separate branch categories below. For the remaining states, the sources offer conflicting guidance on typology, so they remain unclassified.

Table 8: Typology of State Revenue Estimation Practices.

Consensus State	**Executive Branch**	**Separate Branches**	**Unclassified**
Delaware	Arkansas	Alabama	Alaska
Florida	Georgia	Arizona	Connecticut
Indiana	Minnesota	California	Hawaii
Iowa	Oregon	Colorado	Louisiana
Kansas	Texas	Idaho	Mississippi
Kentucky	West Virginia	Illinois	Nevada
Maine		Montana	North Carolina
Maryland		New Hampshire	North Dakota
Massachusetts		New Jersey	Ohio
Michigan		Pennsylvania	Oklahoma
Missouri		Wisconsin	Utah
Nebraska			Virginia
New Mexico			Washington
New York			Wyoming
Rhode Island			
South Carolina			
Tennessee			
Vermont			

Sources:

George A. Krause and James W. Douglas, *Organizational Structure and the Optimal Design of Policymaking Panels: Evidence from Consensus Group Commissions' Revenue Forecasts in the American States*, 57 Am. J. of Pol. Sci. 135, 143 (2013).

NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS, *Budget Processes in the States*, (Spring 2015), 31-38, https://higherlogicdownload.s3.amazonaws.com/NASBO/9d2d2db1-c943-4f1b-b750-0fca152d64c2/UploadedImages/Reports/2015%20Budget%20Processes%20-%20S.pdf.

Elizabeth C. McNichol, *Improving State Revenue Forecasting: Best Practices for a More Trusted and Reliable Revenue Estimate*, CENTER ON BUDGET AND POLICY PRIORITIES 1-3 (Sep. 4, 2014) https://www.cbpp.org/sites/default/files/atoms/files/8-7-14sfp.pdf.

Appendix B: Forecast Error Tables

The following tables include state-level data on forecaster error during the Great Recession, as self-reported from *The Fiscal Survey of the States* compiled by the National Association of State Budget Officers (NASBO).

Table 9: Mean Percentage Error Annual Averages, 2008-2011

	Average Sales MPE	**Average Personal Income Tax MPE**	**Average Corporate Income Tax MPE**
Alabama	6.72%	9.06%	20.83%
Alaska	N/A	N/A	2.75%
Arizona	10.26%	12.84%	20.82%
Arkansas	3.69%	0.89%	-5.60%
California	3.75%	5.43%	1.09%
Colorado	4.41%	1.62%	1.80%
Connecticut	3.06%	3.44%	15.58%
Delaware	N/A	5.75%	-24.71%
Florida	6.69%	N/A	6.82%
Georgia	4.84%	7.28%	-3.40%
Hawaii	1.99%	4.69%	24.02%
Idaho	4.35%	0.93%	19.67%
Illinois	3.41%	2.15%	1.73%
Indiana	3.96%	6.19%	18.46%
Iowa	-2.17%	-14.19%	-1.78%
Kansas	1.09%	2.55%	1.07%
Kentucky	3.91%	3.89%	81.04%
Louisiana	2.40%	-4.21%	31.86%
Maine	0.76%	-0.24%	-2.93%
Maryland	3.59%	4.00%	2.27%
Massachusetts	3.70%	3.17%	-2.28%
Michigan	2.27%	3.83%	10.97%
Minnesota	1.31%	4.05%	5.61%
Mississippi	3.71%	3.42%	5.21%
Missouri	4.59%	6.89%	18.90%
Montana	-2.45%	6.97%	16.23%

Nebraska	0.67%	1.02%	4.27%
Nevada	13.05%	N/A	N/A
New Hampshire	N/A	N/A	8.70%
New Jersey	3.96%	2.02%	1.69%
New Mexico	6.09%	8.68%	46.05%
New York	2.40%	4.04%	7.29%
North Carolina	3.36%	6.19%	4.54%
North Dakota	-6.94%	-7.76%	-3.93%
Ohio	2.52%	2.78%	6.10%
Oklahoma	4.27%	7.52%	42.21%
Oregon	N/A	8.21%	16.79%
Pennsylvania	2.73%	2.94%	6.53%
Rhode Island	2.75%	3.89%	8.29%
South Carolina	5.98%	6.66%	14.37%
South Dakota	0.87%	N/A	N/A
Tennessee	4.36%	5.29%	7.91%
Texas	0.36%	N/A	N/A
Utah	5.82%	3.68%	11.39%
Vermont	3.87%	1.06%	-6.99%
Virginia	4.22%	4.29%	4.26%
Washington	6.94%	N/A	N/A
West Virginia	3.71%	-1.10%	1.99%
Wisconsin	4.54%	3.95%	6.34%
Wyoming	-3.55%	N/A	N/A
Mean MPE	3.30%	3.42%	10.09%
Standard Deviation	0.03	0.05	0.17
75th Percentile	4.39%	6.08%	16.23%
Median MPE	3.70%	3.89%	6.34%
25th Percentile	2.30%	2.05%	1.73%
n	318	291	313

Source: NASBO AND NATIONAL GOVERNORS ASSOCIATION, *The Fiscal Survey of States*, Spring 2007-Spring 2011, https://www.nasbo.org/mainsite/reports-data/fiscal-survey-of-states/fiscal-survey-archives.

Table 10: Mean Absolute Percentage Error (MAPE) and Mean Percentage Error (MPE) 2007-2011

	MPE Average	MAPE Average	(1) Naïve Lag MAPE Average	(2) Naïve Trend MAPE Average	(3) Naïve Lag MAPE - State MAPE	(4) Naïve Trend MAPE - State MAPE
Alabama	12.44%	28.71%	20.27%	23.11%	-8.44%	-5.61%
Alaska	-2.83%	20.40%	14.57%	17.46%	-5.84%	-2.94%
Arizona	15.90%	20.69%	15.17%	28.69%	-5.52%	8.00%
Arkansas	-3.93%	9.26%	8.13%	6.54%	-1.14%	-2.72%
California	0.40%	8.02%	19.56%	9.45%	11.54%	1.43%
Colorado	1.31%	13.12%	11.11%	16.68%	-2.01%	3.56%
Connecticut	8.69%	15.54%	9.96%	10.49%	-5.58%	-5.05%
Delaware	-13.34%	33.55%	29.59%	34.88%	-3.96%	1.33%
Florida	4.50%	12.82%	6.80%	16.69%	-6.02%	3.87%
Georgia	-6.37%	10.08%	12.28%	14.44%	2.20%	4.36%
Hawaii	16.34%	29.57%	22.52%	32.11%	-7.05%	2.54%
Idaho	13.43%	17.17%	16.23%	17.36%	-0.94%	0.19%
Illinois	1.19%	11.46%	16.75%	16.57%	5.30%	5.11%
Indiana	10.32%	19.17%	13.06%	33.38%	-6.11%	14.21%
Iowa	-7.27%	13.25%	10.96%	11.95%	-2.29%	-1.30%
Kansas	-1.76%	7.71%	15.24%	15.78%	7.53%	8.07%
Kentucky	60.09%	65.97%	24.80%	95.90%	-41.17%	29.92%
Louisiana	13.74%	46.59%	43.30%	40.83%	-3.28%	-5.76%
Maine	-4.73%	13.01%	9.29%	13.36%	-3.72%	0.36%
Maryland	5.74%	13.56%	8.79%	19.33%	-4.77%	5.77%
Massachusetts	-5.11%	8.86%	18.20%	11.78%	9.35%	2.92%
Michigan	9.09%	9.09%	9.22%	14.19%	0.12%	5.10%
Minnesota	1.62%	25.32%	20.16%	21.14%	-5.16%	-4.18%
Mississippi	-0.02%	12.49%	7.73%	12.27%	-4.77%	-0.22%
Missouri	12.44%	20.10%	8.69%	15.85%	-11.41%	-4.24%
Montana	3.45%	22.99%	19.42%	18.54%	-3.58%	-4.45%
Nebraska	2.45%	4.81%	10.90%	12.35%	6.08%	7.54%
New Hampshire	4.95%	10.19%	13.19%	10.86%	3.00%	0.67%
New Jersey	-1.10%	8.64%	9.05%	22.83%	0.41%	14.19%
New Mexico	31.07%	43.11%	25.01%	35.93%	-18.09%	-7.18%
New York	2.34%	9.76%	5.31%	9.71%	-4.45%	-0.05%
North Carolina	-2.19%	18.70%	18.02%	8.59%	-0.68%	-10.11%
North Dakota	-15.36%	30.86%	23.29%	27.51%	-7.57%	-3.35%

Ohio	1.18%	13.77%	64.53%	91.79%	50.75%	78.02%
Oklahoma	20.79%	48.55%	27.04%	43.88%	-21.52%	-4.67%
Oregon	5.07%	26.86%	19.43%	25.77%	-7.43%	-1.09%
Pennsylvania	3.41%	6.83%	6.10%	9.17%	-0.73%	2.33%
Rhode Island	5.86%	13.95%	15.85%	15.20%	1.90%	1.25%
South Carolina	7.30%	24.07%	34.41%	22.43%	10.34%	-1.64%
Tennessee	5.54%	11.03%	7.42%	10.09%	-3.61%	-0.93%
Utah	0.59%	21.03%	17.30%	19.69%	-3.73%	-1.34%
Vermont	-10.48%	22.59%	28.36%	13.92%	5.77%	-8.67%
Virginia	0.53%	6.89%	6.27%	6.21%	-0.63%	-0.68%
West Virginia	-0.50%	6.35%	13.71%	12.68%	7.36%	6.33%
Wisconsin	-0.47%	17.23%	11.81%	9.62%	-5.42%	-7.61%
National Average	**4.58%**	**18.97%**	**17.08%**	**21.71%**	**-1.89%**	**2.74%**

Source: NASBO AND NATIONAL GOVERNORS ASSOCIATION, *The Fiscal Survey of States*, Spring 2007-Spring 2011, https://www.nasbo.org/mainsite/reports-data/fiscal-survey-of-states/fiscal-survey-archives.

Bibliography

Donald J. Boyd & Lucy Dadayan, *State Tax Revenue Forecasting Accuracy Technical Report*, THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT (Sep. 2014).

Richard T. Boylan, *Political Distortions in State Forecasts,* 136 PUB. CHOICE 411 (2008).

[1]P. Delp *et al.*, *Nominal Group Technique,* SYSTEM TOOLS FOR PROJECT PLANNING, International Development Institute, http://www.aucd.net/docs/urc/Leadership_Institute/Subsequent%20Leadership%20Institute%20Materials/Nominal%20Group%20Technique.pdf.

Elaine Deschamps, *The impact of institutional change on forecast accuracy: A case study of budget forecasting in Washington State,* 20 INT'L J. OF FORECASTING 647 (2004).

John S. Dryzek, *Legitimacy and Economy in Deliberative Democracy,* 29 POL. THEORY 651 (2001).

Daniel Gigone & Reid Hastie, *Proper Analysis of the Accuracy of Group Judgments,* 121 PSYCH. BULLETIN 149 (1997).

Reid Hastie & Tatsuya Kameda, *The Robust Beauty of Majority Rules in Group Decisions,* 112 PSYCHOLOGICAL REVIEW 494 (2005).

Jack Healey, *Built Up By Oil Boom, North Dakota Now Has an Emptier Feeling*, N.Y. TIMES (Feb. 7, 2016) https://www.nytimes.com/2016/02/08/us/built-up-by-oil-boom-north-dakota-now-has-an-emptier-feeling.html.

Chia-Chien Hsu and Brian A. Sandford, *The Delphi Technique: Making Sense of Consensus,* 12 PRACTICAL RESEARCH AND ASSESSMENT 1 (2007).

Richard F. Keevey, *Budget Basics: How the State Budget is Developed and Revenue Estimated*, NJ SPOTLIGHT (Oct. 13, 2017) http://www.njspotlight.com/stories/17/10/12/budget-basics-how-the-state-budget-is-developed-and-revenue-estimated/.

George A. Krause & James W. Douglas, *Organizational Structure and the Optimal Design of Policymaking Panels: Evidence from Consensus Group Commissions' Revenue Forecasts in the American States,* 57 AM. J. OF POL. SCI. 135 (2013).

George A. Krause, David E. Lewis, & James W. Douglas, *Political Appointments, Civil Service Systems, and Bureaucratic Competence: Organizational Balancing and Executive Branch Revenue Forecasts in the American States,* 50 AM. J. POL. SCI. 770 (2006).

Barbara Luppi & Francesco Parisi, *Jury Size and the Hung-Jury Paradox,* 42 JOURNAL OF LEGAL STUDIES 2, 399 (June 2013).

Robert C. Luskin, James S. Fishkin, & Roger Jowell, *Considered Opinions: Deliberative Polling in Britain*, 32 BRITISH J. OF POL. SCI. 455 (2002).

Elizabeth C. McNichol, *Improving State Revenue Forecasting: Best Practices for a More Trusted and Reliable Revenue Estimate,* CENTER. ON BUDGET AND POLICY PRIORITIES 1 (Sep. 4, 2014) https://www.cbpp.org/sites/default/files/atoms/files/8-7-14sfp.pdf.

John L. Mikesell & Justin M. Ross, *State Revenue Forecasts and Political Acceptance: The Value of Consensus Forecasting in the Budget Process,* 74 PUB. ADMIN. REV. 188 (2014).

NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS & THE NATIONAL GOVERNORS ASSOCIATION, *The Fiscal Survey of States* (Fall 2008-Spring 2011) https://www.nasbo.org/mainsite/reports-data/fiscal-survey-of-states/fiscal-survey-archives.

NATIONAL ASSOCIATION OF STATE BUDGET OFFICERS, BUDGET PROCESSES IN THE STATES (Spring 2015), https://higherlogicdownload.s3.amazonaws.com/NASBO/9d2d2db1-c943-4f1b-b750-0fca152d64c2/UploadedImages/Reports/2015%20Budget%20Processes%20-%20S.pdf.

THE PEW CENTER ON THE STATES AND THE NELSON A. ROCKEFELLER INSTITUTE OF GOVERNMENT, *States' Revenue Estimating: Cracks in the Crystal Ball* (March 2011), http://www.rockinst.org/pdf/government_finance/2011-03-01-states-revenue-estimating-report.pdf.

James M. Poterba, *State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Policies,* 102 J. OF POL. ECON. 799, 803 (1994).

Shanna Rose & Daniel L. Smith, *Budget Slack, Institutions and Transparency,* 72 PUB. ADMIN. R. 187 (2011).

Gene Rowe & George Wright, *Expert Opinions in Forecasting: The Role of the Delphi Technique, in* PRINCIPLES OF FORECASTING, 125 (J.S. Armstrong, ed., 2001).

Garold Stasser *et al., Information Sampling in Structured and Unstructured Discussions of Three- and Six-Person Groups,* 57 J. PERSONALITY & SOC. PSYCH. 67 (1989).

Garold Stasser & Beth Dietz-Uhler, *Collective Choice Judgment, and Problem Solving*, *in* BLACKWELL HANDBOOK OF GROUP PSYCHOLOGY: GROUP PROCESSES (Michael A. Hogg and R. Scott Tindale, eds., 2001).

Cass R. Sunstein. *Group Judgments: Statistical Means, Deliberation, and Information Markets*, 80 N.Y.U. L. REV. 962 (2005).

CASS R. SUNSTEIN & REID HASTIE, WISER: GETTING BEYOND GROUPTHINK TO MAKE GROUPS SMARTER (2015).

Jeffrey M. Tebbs, *Breaking the Stalemate: A Proposal for a Consensus Revenue Forecasting Process*, CONNECTICUT VOICES FOR CHILDREN http://www.ctvoices.org/sites/default/files/bud09revenueforecasting.pdf (March 2009).

TEX. CONST. Art. III, Sec. 49a.

Glen Whyte & James K. Sebenius, *The Effect of Multiple Anchors on Anchoring in Individual and Group Judgment*, 69 ORG. BEH. AND HUMAN DECISION PROCESSES 75 (January 1997)

[1]Daniel W. Williams & Thad D. Calabrese, *The Status of Budget Forecasting*, 2 J. OF PUB. AND NONPROFIT AFFAIRS 127 (2016).